

(19) 世界知的所有権機関
国際事務局(43) 国際公開日
2005 年 10 月 13 日 (13.10.2005)

PCT

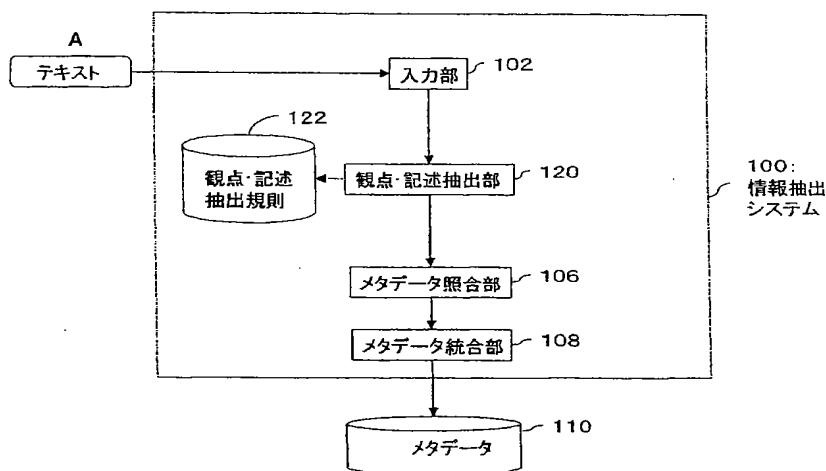
(10) 国際公開番号
WO 2005/096182 A1

- (51) 国際特許分類⁷: G06F 17/30 (72) 発明者; および
(21) 国際出願番号: PCT/JP2005/006317 (75) 発明者/出願人 (米国についてのみ): 野本 昌子
(22) 国際出願日: 2005 年 3 月 31 日 (31.03.2005) (74) 代理人: 蔵合 正博, 外 (ZOGO, Masahiro et al.); 〒
(25) 国際出願の言語: 日本語 1020083 東京都千代田区麹町 5 丁目 7 番地 秀和紀尾
(26) 国際公開の言語: 日本語 井町 T B R ビル Tokyo (JP).
(30) 優先権データ: 特願 2004-107139 2004 年 3 月 31 日 (31.03.2004) JP (81) 指定国 (表示のない限り、全ての種類の国内保護が
可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,
BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA,
NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE,
SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, YU, ZA, ZM, ZW.

[続葉有]

(54) Title: INFORMATION EXTRACTION SYSTEM

(54) 発明の名称: 情報抽出システム



A... TEXT
102... INPUT UNIT
122... VIEWPOINT/DESCRIPTION EXTRACTION RULE
120... VIEWPOINT/DESCRIPTION EXTRACTION UNIT
106... META DATA CORRELATION UNIT
108... META DATA UNIFICATION UNIT
110... META DATA
100... INFORMATION EXTRACTION SYSTEM

(57) Abstract: It is possible to extract a description content as a fact and an opinion on a matter expressed in a text while correlating them with the viewpoint and description of the fact and the opinion. An information extraction system includes an input unit for inputting a text; a viewpoint/description extraction rule storage unit containing a viewpoint/description extraction rule for identifying a pair of viewpoint of expression described in the text and a description on the viewpoint; a viewpoint/description extraction unit for correlating a viewpoint and its description in a pair by using the viewpoint/description extraction rule from at least one of the attributes: a syntactic attribute or lexical attribute and extracting the pair as element meta data to which identification information for identifying them is attached; and a meta data storage unit for storing element meta data extracted by the viewpoint/description extraction unit. The description content such as a fact and an opinion on a matter expressed in a text is formed as a

pair of viewpoint and description and extracted in a format in which facts and opinions are correlated.

[続葉有]



(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

(57) 要約:【課題】 テキスト中に表現された事物に関する事実や意見などの記述内容を、事実や意見の観点と記述を対応付けて抽出する。そのために、テキストを入力する入力部と、テキストに記述された表現の観点とその観点に関する記述の組を特定するための観点・記述抽出規則を格納する観点・記述抽出規則格納部と、前記テキスト中の文字列の統語的属性または意味的属性の少なくとも一方の属性から、前記観点・記述抽出規則を用いて観点とその記述の組を対応付け、それらを識別するための識別情報を付与した要素メタデータとして抽出する観点・記述抽出部と、前記観点・記述抽出部が抽出した要素メタデータを格納するメタデータ格納部を有する構成とし、テキスト中に表現された事物に関する事実や意見などの記述内容を、観点と記述の組として構成し、事実や意見の対応付けた形に整理して抽出する。